A system for Rapid Analysis of Transactional Insurance Data to identify Trends in Cost of Work-Related Injuries

Alan Ducatman, M.D,M.Sc, Syed Islam, MBBS, Dr.PH, John Atkins, Ph.D, Edward Doyle, M.D, M.Sc, James Harner, Ph.D

Nature of Insurance Data and Utility of Data Cubes

Insurance data are stored in multiple independent tables and contain one or more outcome variables (facts) of interest, such as financial transactions related to medical costs, indemnity costs and frequency of specific diagnosis or procedures. Analysis can be ponderous, if these facts have to be stratified by various dimensions such as employer, date of accident, diagnosis (ICD-9 groups), service provider, services rendered (CPT codes), location (of patient, injury, or provider) etc. as they are stored in many different tables that do not have easy linkage. One way to rapidly analyze such data is to pre-aggregate an outcome variable (fact) across various dimensions (variables that we want to compare), creating a multi-dimensional 'data cube'. Therefore, each cube has dimensions that are distinct and chosen for the ability to influence decision making. A single cube may contain one or more facts based on feasibility of pre-aggregation by various dimensions.

What is needed for analysis of large multidimensional cubes?

- For rapid analysis of large datasets, we need specific software or a database <u>frontend</u>.

- A database frontend presents the data in a way that the user can understand.

- It is intuitive
  - It has a friendly user interface.
  - It requires no database knowledge to use.
  - It allows the users to explore data easily.

What is Web-based executive information system (WebEIS)?

- WebEIS is a web-based <u>frontend</u> that can rapidly analyze multidimensional data cubes.

- The WebEIS used in this presentation was developed by SAS Institute.

- As the name suggests, WebEIS is web based and requires only a web browser.

Why web based?

- Requires no special software to run
  - No hassle of setting up and maintaining software
  - Little training needed for the users
- Provides easy access
  - Users can logon anywhere through the Internet.
  - It is as easy to load as a web page.
  - Users have indirect access to the server machine, making it more secure.

Other features

- Security features are built in to deny unwanted access.

Schematic for Health Care Data Standardization Components
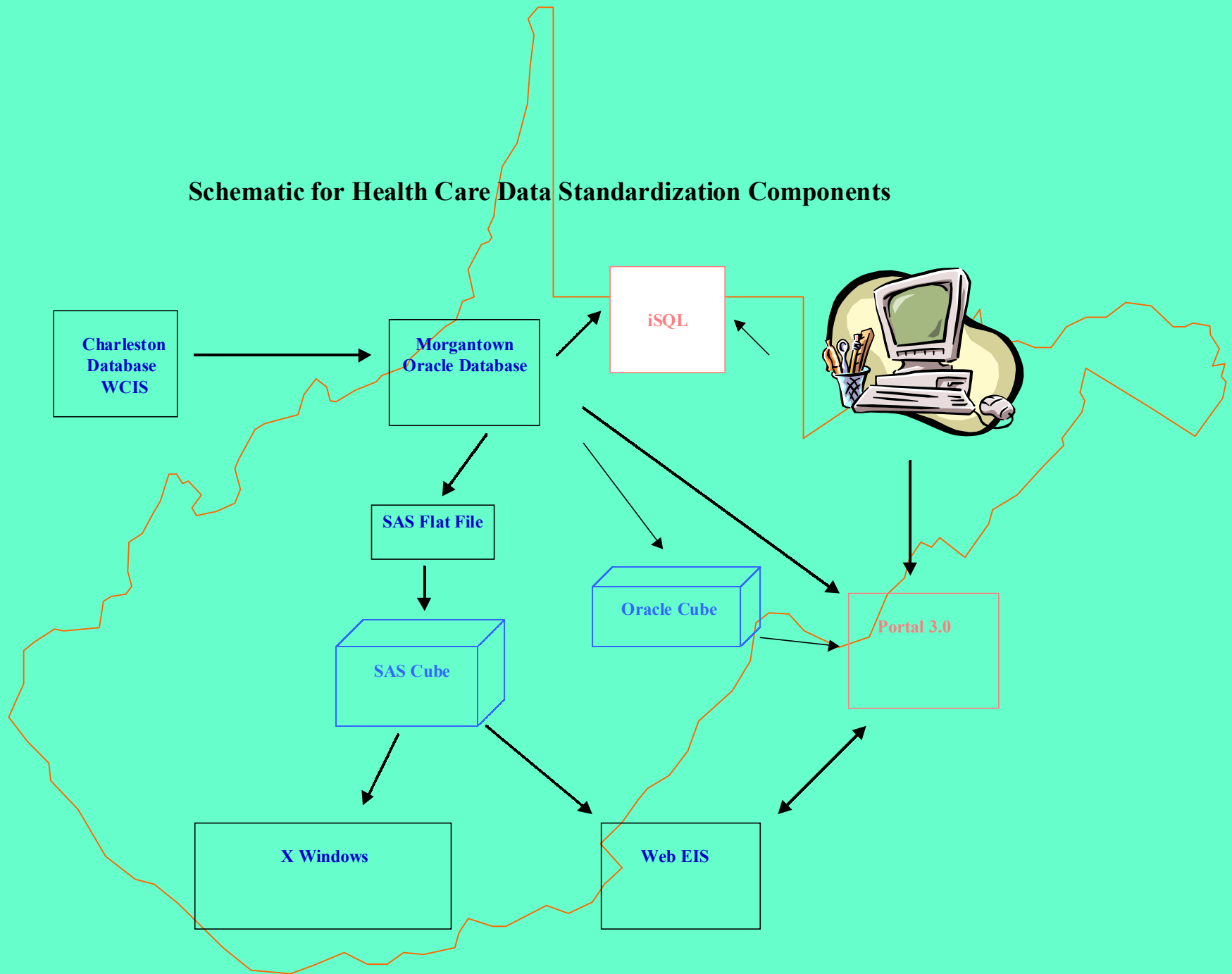
West Virginia Workers' Compensation Data System

Let us suppose that the claim managers want to know the average cost of physical therapy *(A)* for back injuries *(B)* of 1999 *(C)* and want to know the location of accident *(D)* and type of providers *(B)*

A. Medical Bill Detail
Every Medical
Transaction
Is captured
Unique ID-" Bill control
Number"
**No Claim number
Available**
*Approx. 30 million
records*

B. Medical Bill Summary
Unique ID-" Bill
control number"
**Claim number
available approx.**
12 million records

C. Claim table
Unique ID- Claim
number
many claim related
Information available
1.5 million records

D. Accident table
Unique ID
Claim number
1.5 million record

A.
This table has all medical costs and associated with procedure codes including medications and providers

B.
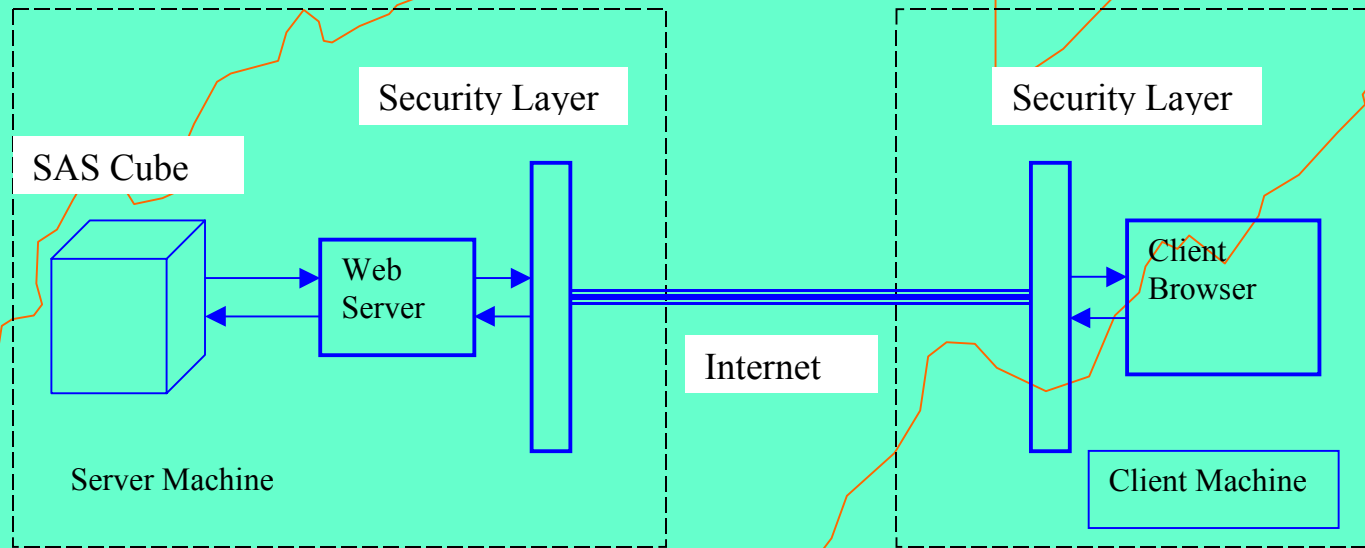This table has diagnosis codes associated with each claim

C.
Various claim characteristics and accident date

D.
Exposure information and accident location

For the above question we need to merge A+B+C+D

How does it work?

Data Visualization

- WebEIS can organize aggregated data into tabular report or charts. To produce a table, users only need to drag and drop variables on the rows or columns of the table. To produce a chart, they drag and drop variables on the x-axis and y-axis.

- Let's look at this example.

cost_by_year - Microsoft Internet Explorer

File   Edit   View   Favorites   Tools   Help

←Back ▼ → ▼ ⊗ ▣ ⌂ | ◎Search ▣Favorites ◐History | ▣▼ 🖨 ⊞ ▼

Address 🔲 C:\AppDevStudio\webEIS\workcomp\cost_by_year.html   ▼ ⌀Go

- ◆ cost_by_year
  - ◆ Tabular report
  - ◆ chart

year: 1992, 1993, 1994, 1995, 1996, 1997, 1998

aregion	Region 1		Region 2		Region 3	
	SERVICE_ALLOW_AMT		SERVICE_ALLOW_AMT		SERVICE_ALLOW_AMT	
year	Sum	Percent of Sum	Sum	Percent of Sum	Sum	Percent of Sum
1992	$22,828,194	14.17	$14,049,493	15.18	$8,519,850	10.16
1993	$21,769,428	13.51	$15,450,119	16.69	$10,027,162	11.96
1994	$24,786,755	15.39	$16,771,840	18.12	$13,249,422	15.80
1995	$21,811,513	13.54	$10,888,820	11.76	$16,469,104	19.64
1996	$22,797,175	14.15	$11,258,254	12.16	$11,802,494	14.08
1997	$25,447,895	15.80	$12,419,060	13.42	$12,260,850	14.62
1998	$21,636,985	13.43	$11,735,905	12.68	$11,514,382	13.73

📄 Contents 📦 Data

Applet started.   🖥 My Computer

🏁Start | ▣ 🌐 🗗 | 🗀workco... | 🎚 Start SA... | 📑Presenta... | PowerPoint ... | 🌐cost_b... | 🗄 BATCH S... | 4:30 PM

cost_by_year - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back | → | Search | Favorites | History | W

Address C:\AppDevStudio\webEIS\workcomp\cost_by_year.html | Go

File ∇ View ∇ Bookmarks ∇ Help ∇

- cost_by_year
  - Tabular report
  - chart

year: 1992, 1993, 1994, 1995, 1996, 1997, 1998

aregion			Region 1		Region 2		Region 3	
			SERVICE_ALLOW_AMT		SERVICE_ALLOW_AMT		SERVICE_ALL...	
year	quarter	month	Sum	Percent of Sum	Sum	Percent of Sum	Sum	Pe
1992			$22,828,194	14.17	$14,049,493	15.18	$8,519,850	
1993			$21,769,428	13.51	$15,450,119	16.69	$10,027,162	
1994			$24,786,755	15.39	$16,771,840	18.12	$13,249,422	
1995	Q1		$5,706,765	26.16	$2,591,249	23.80	$8,731,672	
	Q2	Apr	$1,741,486	33.61	$763,537	32.23	$742,956	
		May	$2,018,629	38.96	$786,041	33.18	$939,586	
		Jun	$1,421,276	27.43	$819,320	34.59	$926,019	
	Q3		$5,745,189	26.34	$2,949,184	27.08	$2,590,952	
	Q4		$5,178,168	23.74	$2,979,489	27.36	$2,537,920	
1996			$22,797,175	14.15	$11,258,254	12.16	$11,802,494	
1997			$25,447,895	15.80	$12,419,060	13.42	$12,260,850	
1998			$21,636,985	13.43	$11,735,905	12.68	$11,514,382	

Applet started.   My Computer

Start | Sta... | wo... | dia... | BA... | Pre... | Power... | cos... | co... | BA... | 4:46 PM

time_region_ind - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back → ⊗ ⟳ ⌂ Search Favorites History

Address C:\AppDevStudio\webEIS\workcomp\time_region_ind.html Go

File ▽ View ▽ Bookmarks ▽ Help ▽

- time_region_ind
  - tabular report
  - chart
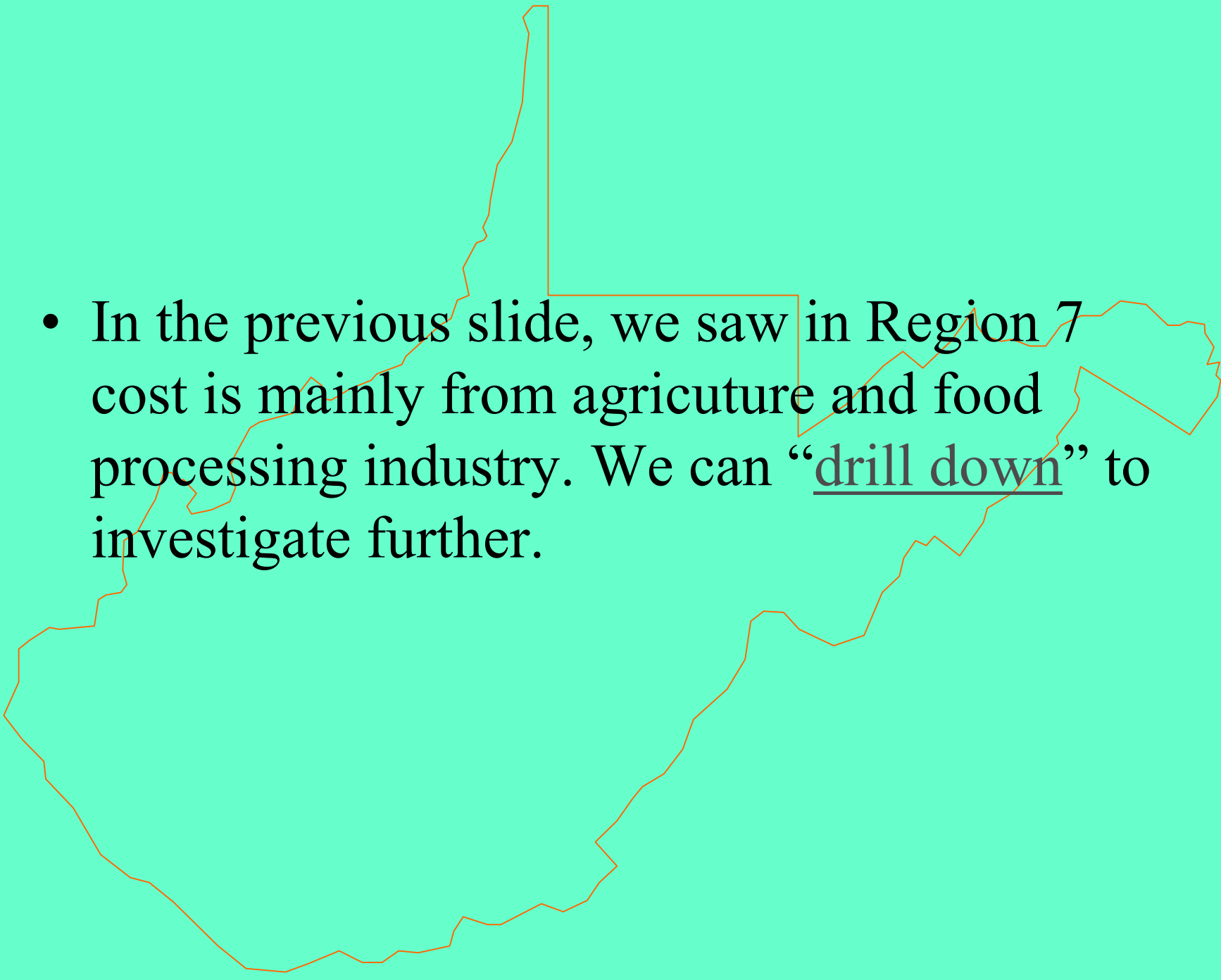
year: 1995, 1998
aregion: Region 1, Region 7
DIAGNOSIS_GROUP: Entrapment neuropathy, upper extremity CTS

year	aregion	INDUSTRY_TEAM	SERVICE_ALLOW_AMT Sum	SERVICE_ALLOW_AMT Percent of Sum
1995	Region 1	Agriculture - Food Processing - Oil & Gas	$19,185	2.39
		Construction	$72,121	8.99
		Education & Glass Manufacturing	$55,203	6.88
		Entertainment & Services 1	$24,686	3.08
		Entertainment & Services 2	$53,305	6.65
		Government & Clerical	$202,694	25.28
		Health Care 1	$67,665	8.44
		Health Care 2	$27,465	3.43
		Manufacturing & Timbering	$24,870	3.10
		Metals - Chemicals - Utilities	$13,108	1.63
		Mining 1	$58,617	7.31
		Mining 2	$40,518	5.05
		Mining 3	$62,832	7.84
		Retail I	$8,213	1.02
		Retail II	$56,104	7.00
		Transportation	$15,300	1.91
	Region 7	Agriculture - Food Processing - Oil & Gas	$82,081	30.57
		Construction	$24,278	9.04
		Education & Glass Manufacturing	$9,818	3.66

Applet started. My Computer

Start Start ... workc... Prese... Prese... PowerPoin... time_... BATC... 4:48 PM

- Data can be sliced and diced with multiple dimensions, and dimensions can be "pivoted" (re-arranged) to organize data in different ways, illustrated by this example that looks at costs for one of our 21 diagnoses groups, nerve entrapment syndrome.We will look at the costs for this condition as a function of industry groups in two economic regions for two different years.   Note that a feature, such as high cost, can be highlighted.

- In the previous slide, we saw in Region 7 cost is mainly from agricuture and food processing industry. We can "drill down" to investigate further.
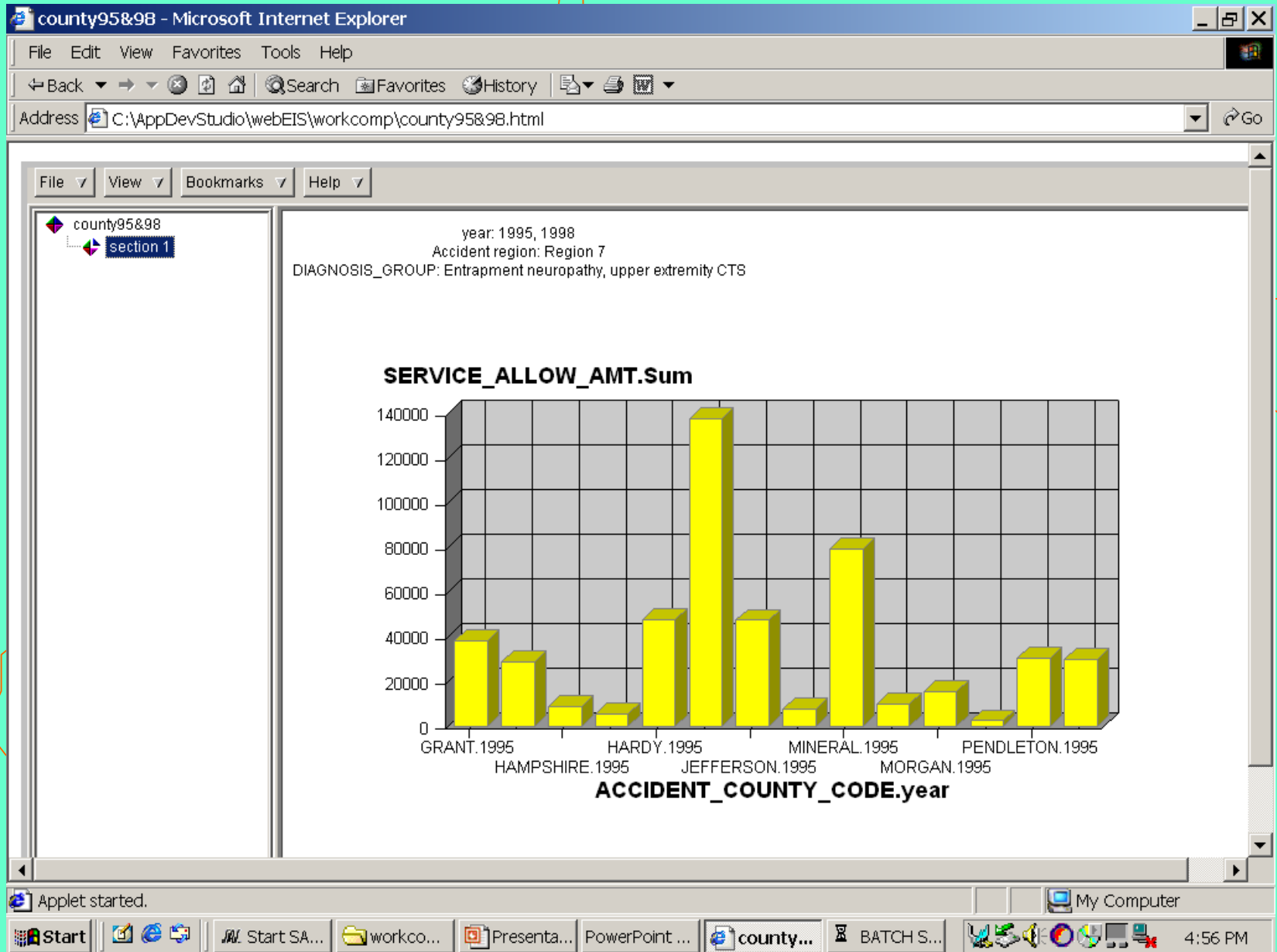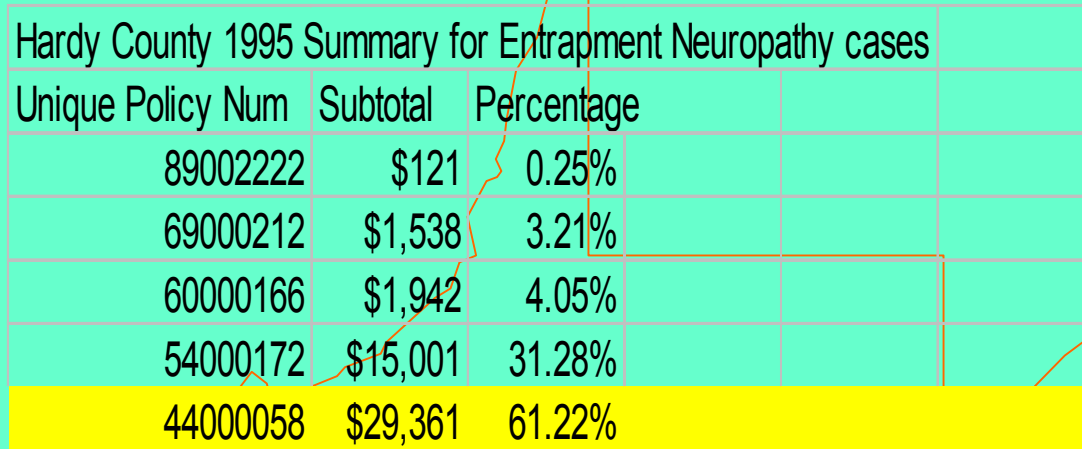
Detail data for region 7

Detail Data ☒
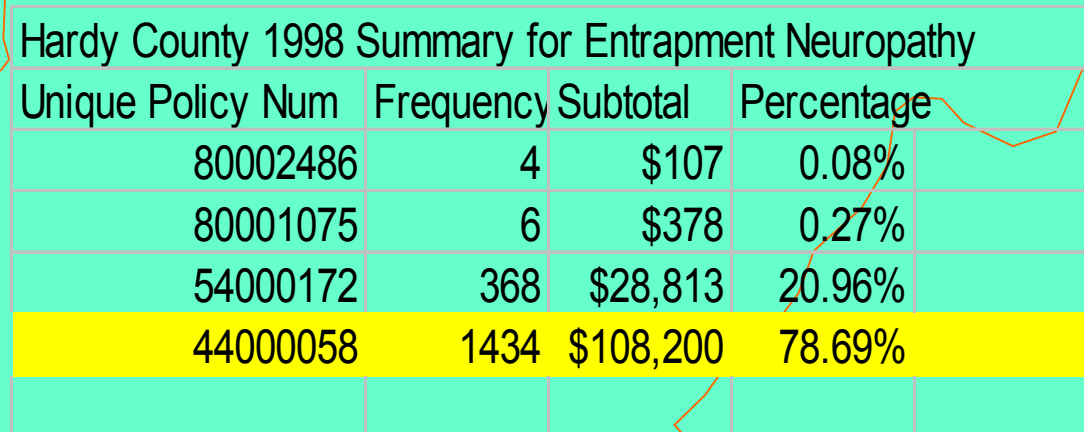
7

	cptclass	save	diagnose	temp	DIAGNOSIS_CODE	ACCIDENT_COUNTY_C
1	unknown	.	sprain	sprain	840.9	HARDY
2	Admin	4435.1	sprain	sprain	840.9	HARDY
3	Admin	1170	sprain	sprain	840.9	HARDY
4	95861-TC	34.62	sprain	sprain	840.9	HARDY
5	95860-26	79.64	sprain	sprain	840.9	HARDY
6	95900-26	71.84	sprain	sprain	840.9	HARDY
7	95904-26	64.05	sprain	sprain	840.9	HARDY
8	Outpt	25.74	sprain	sprain	840.9	HARDY
9	unknown	.	sprain	sprain	840.9	HARDY
10	unknown	.	Trauma/Wound	Trauma/Wound	883.0	HARDY
11	Admin	886.36	Trauma/Wound	Trauma/Wound	883.0	HARDY
12	Drug	8.14	Trauma/Wound	Trauma/Wound	883.0	HARDY
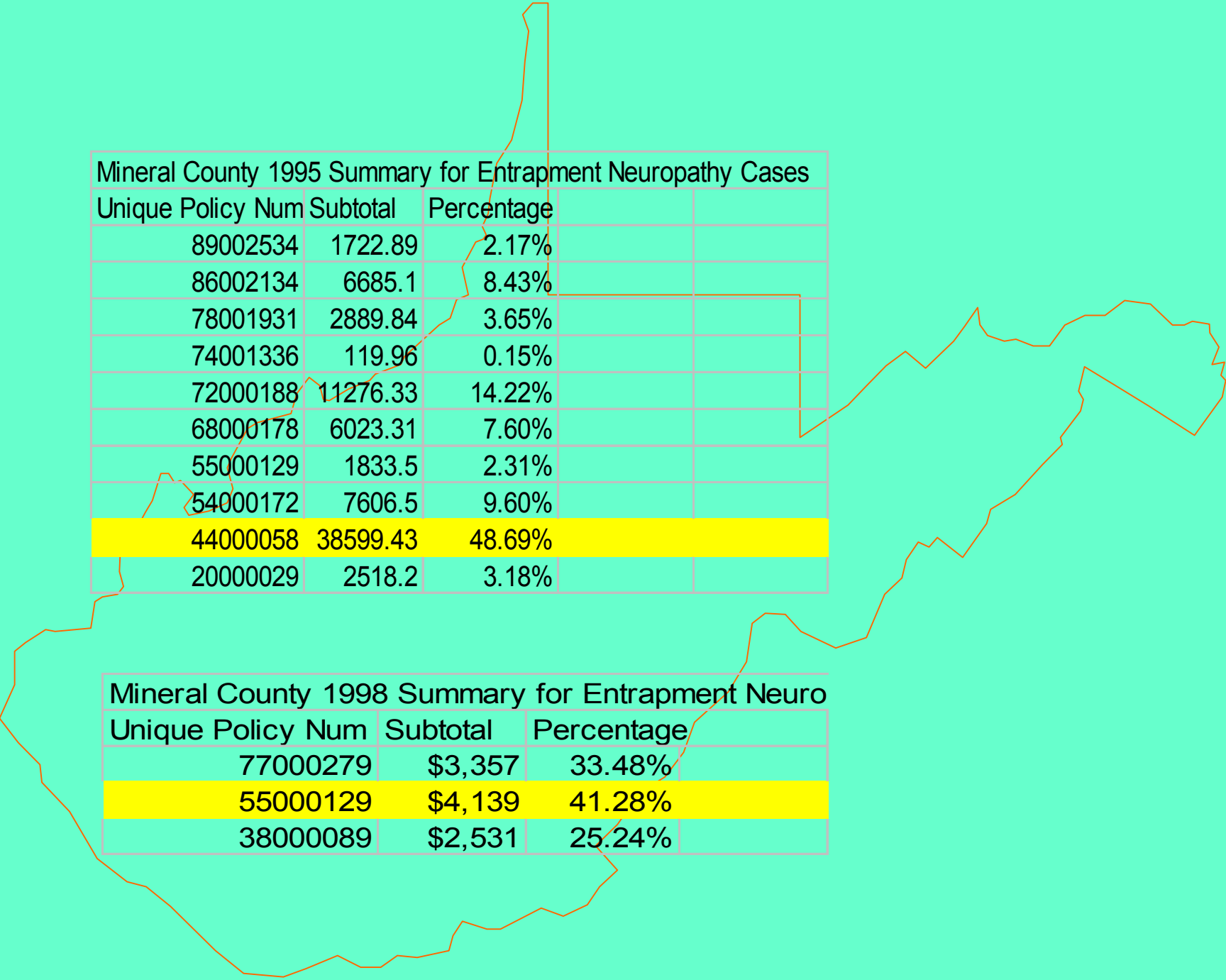
Export to Excel   Close

Hardy County 1995 Summary for Entrapment Neuropathy cases				
Unique Policy Num	Subtotal	Percentage		
89002222	$121	0.25%		
69000212	$1,538	3.21%		
60000166	$1,942	4.05%		
54000172	$15,001	31.28%		
44000058	$29,361	61.22%		

Hardy County 1998 Summary for Entrapment Neuropathy				
Unique Policy Num	Frequency	Subtotal	Percentage	
80002486	4	$107	0.08%	
80001075	6	$378	0.27%	
54000172	368	$28,813	20.96%	
44000058	1434	$108,200	78.69%	

Mineral County 1995 Summary for Entrapment Neuropathy Cases

Unique Policy Num	Subtotal	Percentage		
89002534	1722.89	2.17%		
86002134	6685.1	8.43%		
78001931	2889.84	3.65%		
74001336	119.96	0.15%		
72000188	11276.33	14.22%		
68000178	6023.31	7.60%		
55000129	1833.5	2.31%		
54000172	7606.5	9.60%		
44000058	38599.43	48.69%		
20000029	2518.2	3.18%		

Mineral County 1998 Summary for Entrapment Neuro

Unique Policy Num	Subtotal	Percentage	
77000279	$3,357	33.48%	
55000129	$4,139	41.28%	
38000089	$2,531	25.24%	

Screenshot of WebEIS demo page

List of Reports - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://157.182.239.171/~mddb/

Worker's Compensation WebEIS Demo Page

The web based frontend to the WORKCOMP data warehouse is demonstrated in this page. A list of sample reports are listed below. They are built using SAS webEIS. No special software is required to view the reports but you may need to download Java Plug-in. You will also need a username and password because only authorized users are allowed to view the report.

Click on a link below to view the report.

- Amount by time report
- Amount by accident location and time report
- Amount by accident location report
- Amount by home location report
- Amount by age report
- Amount by specialty code report
- Amount by service code report

Bar chart showing total payments for injured year
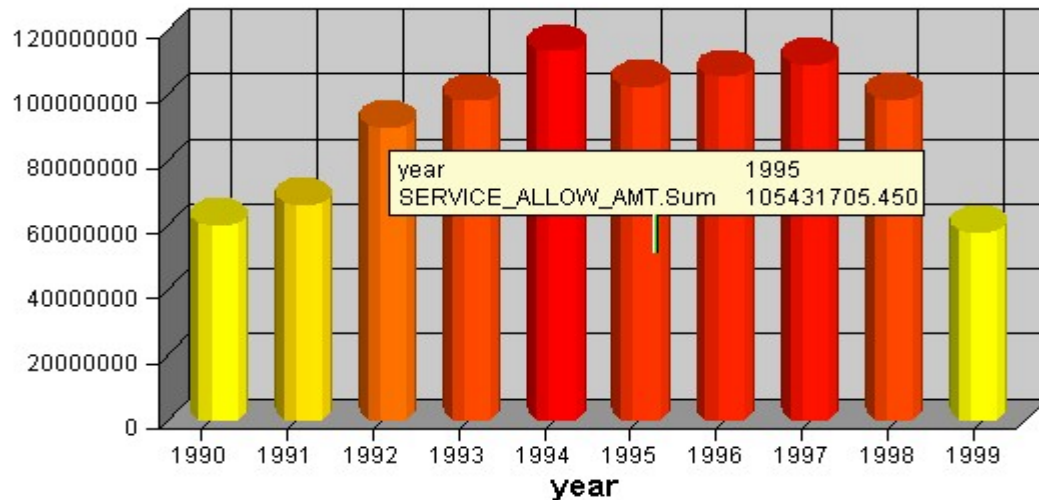
Bar chart showing monthly summary for 1994

Injured Year: 1994
quarter: All

SERVICE_ALLOW_AMT.Sum

SERVICE_ALLOW_AMT.Sum

7582495 — 11969419
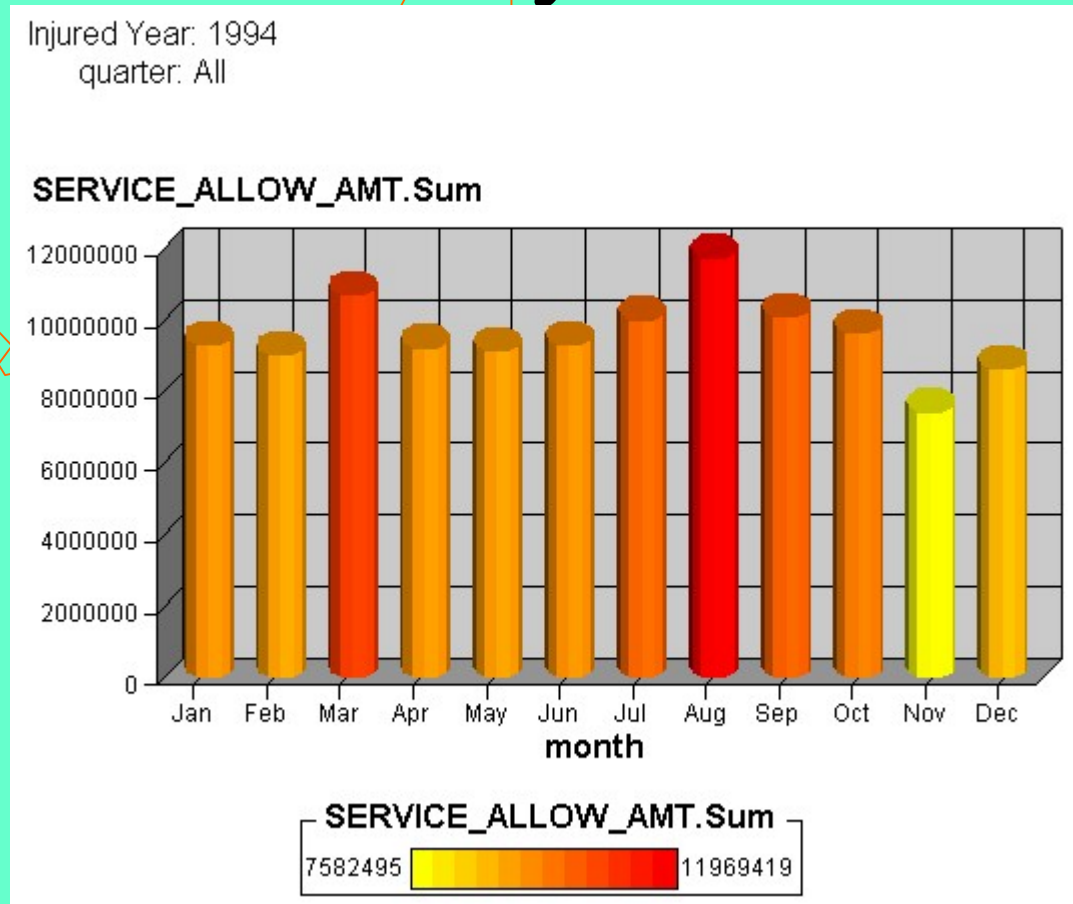
Bar chart showing total payment for different regions

Injured Year: 1994

SERVICE_ALLOW_AMT.Sum

aregion

SERVICE_ALLOW_AMT.Sum

5449910 24786754
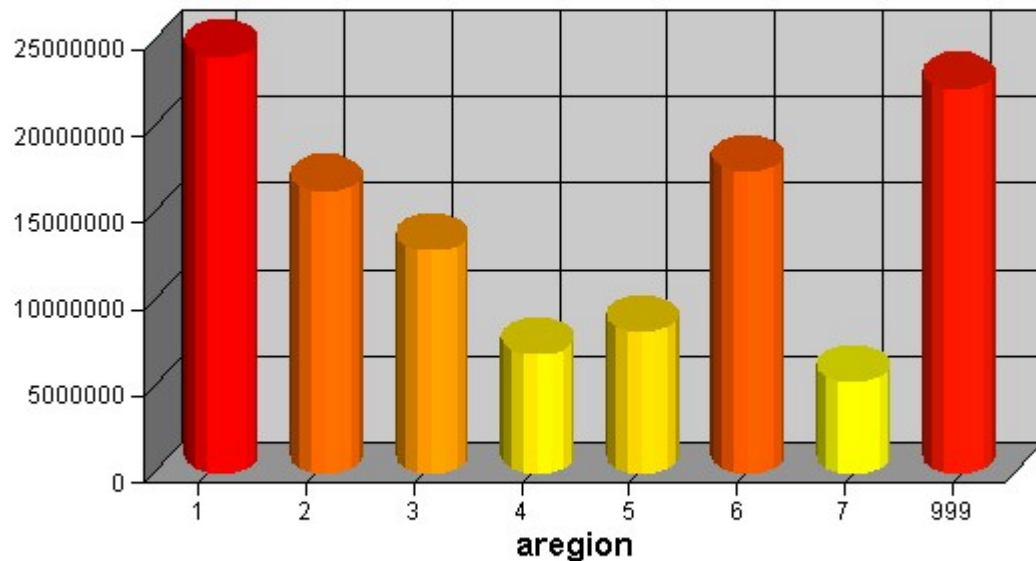
Year by year comparison for the regions
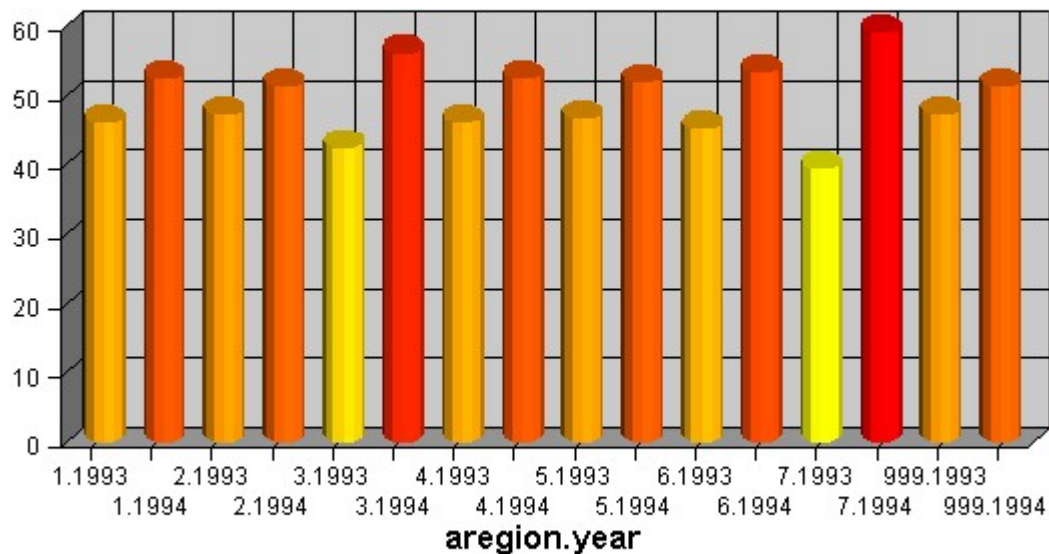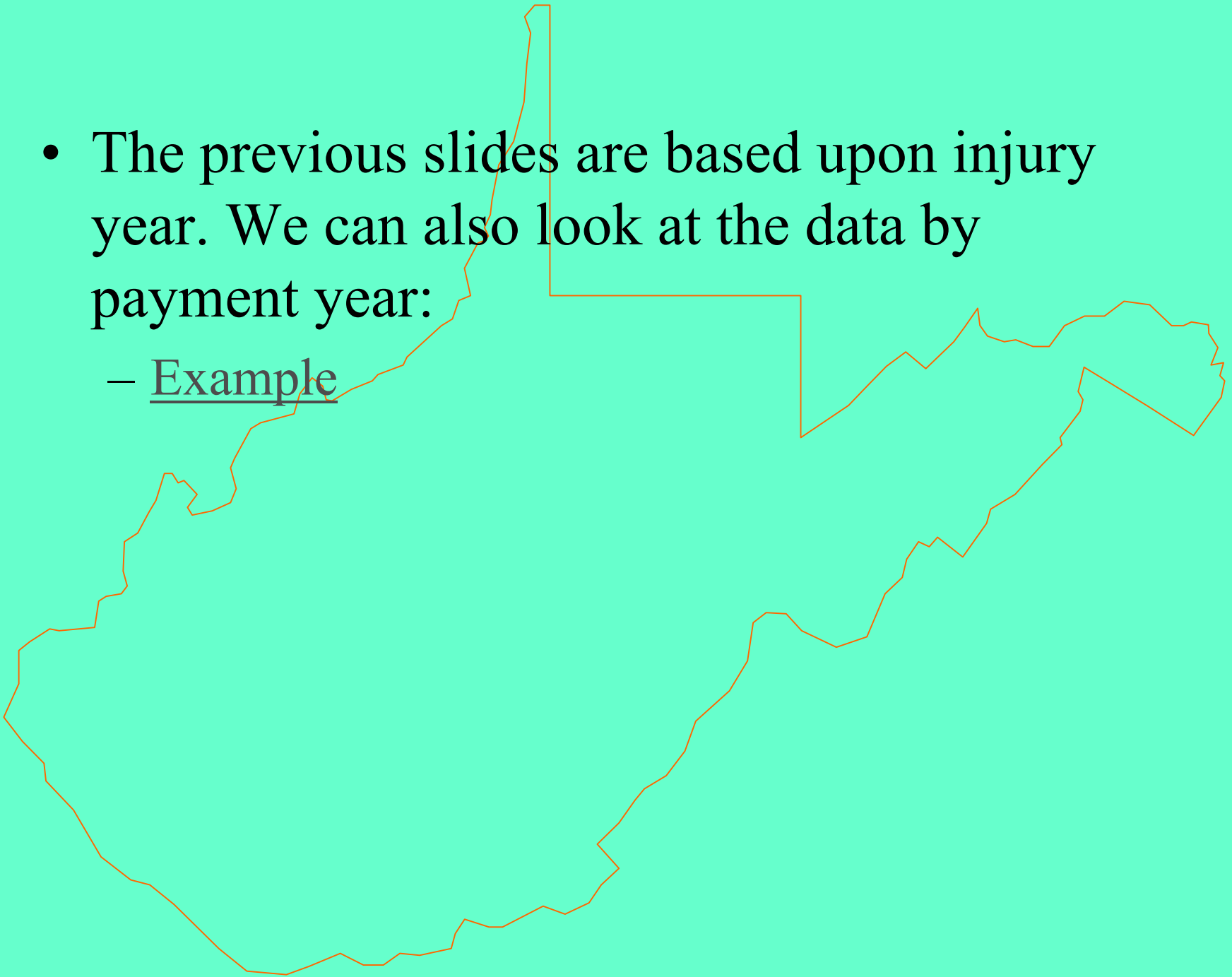
- The previous slides are based upon injury year. We can also look at the data by payment year:
  - Example

diag_spec_ptime - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address C:\AppDevStudio\webEIS\workcomp\diag_spec_ptime.html

diag_spec_ptime
  tabular report

PAYMENT YEAR: 1995, 1999

pyear		1995		1999	
		SERVICE_ALLOW_AMT		SERVICE_ALLOW_AMT	
DIAGNOSIS_GROUP	PROVIDER_CATEGORY	Sum	Percent of Sum	Sum	Percent of Sum
Musculoskeletal (excluding back)	Administrative	$118,921	3.77	$146,077	3.78
	Chiropractor	$102,098	3.24	$106,541	2.76
	Equipment	$18,220	0.58	$47,991	1.24
	Hospital	$1,418,613	45.01	$1,285,499	33.26
	Legal	$8,953	0.28	$21,935	0.57
	NH/Home	$6,144	0.19	$14,693	0.38
	PT/rehab	$317,014	10.06	$912,250	23.60
	Pharmacy	$74,814	2.37	$155,822	4.03
	Physicians	$238,401	7.56	$289,912	7.50
	Radiology	$44,962	1.43	$55,440	1.43
	Surgeons	$615,424	19.53	$770,470	19.93
	Transport	$1,474	0.05	$2,118	0.05
	Unknown	$162,287	5.15	$33,854	0.88
	Vocrehab	$24,423	0.77	$22,377	0.58
Sprain	Administrative	$485,351	3.37	$464,768	3.47
	Chiropractor	$393,167	2.73	$350,378	2.61
	Equipment	$164,485	1.14	$196,265	1.46
	Hearing	$1,250	0.01	$550	0.00

Contents   Data

Applet started.   My Computer

Start   Start SA...   workco...   Microsof...   PowerPoint ...   diag_sp...   BATCH S...   4:43 PM

- To investigate further, we just click on the "Export" button and this will download a comma delimited file for the specific strata which can be read by SAS.

- In SAS, simple analysis can be done to further identify the reasons for anomaly. <u>However, this will require training and experience in using SAS.</u>

- Many questions raised by rating and financial offices will need initial simple SAS analysis. Therefore, training of WC staff on SAS will be extremely useful for further utilization of WebEIS system.

Gender specific *relative risk of injury by age category among cooks over a four-year period

Risk of inju

	14-19	20-29	30-39	40-54	55-64	>=65
Female	0.19	0.31				
Male			0.77	0.21	0.09	0.24

Age category

Female
Male
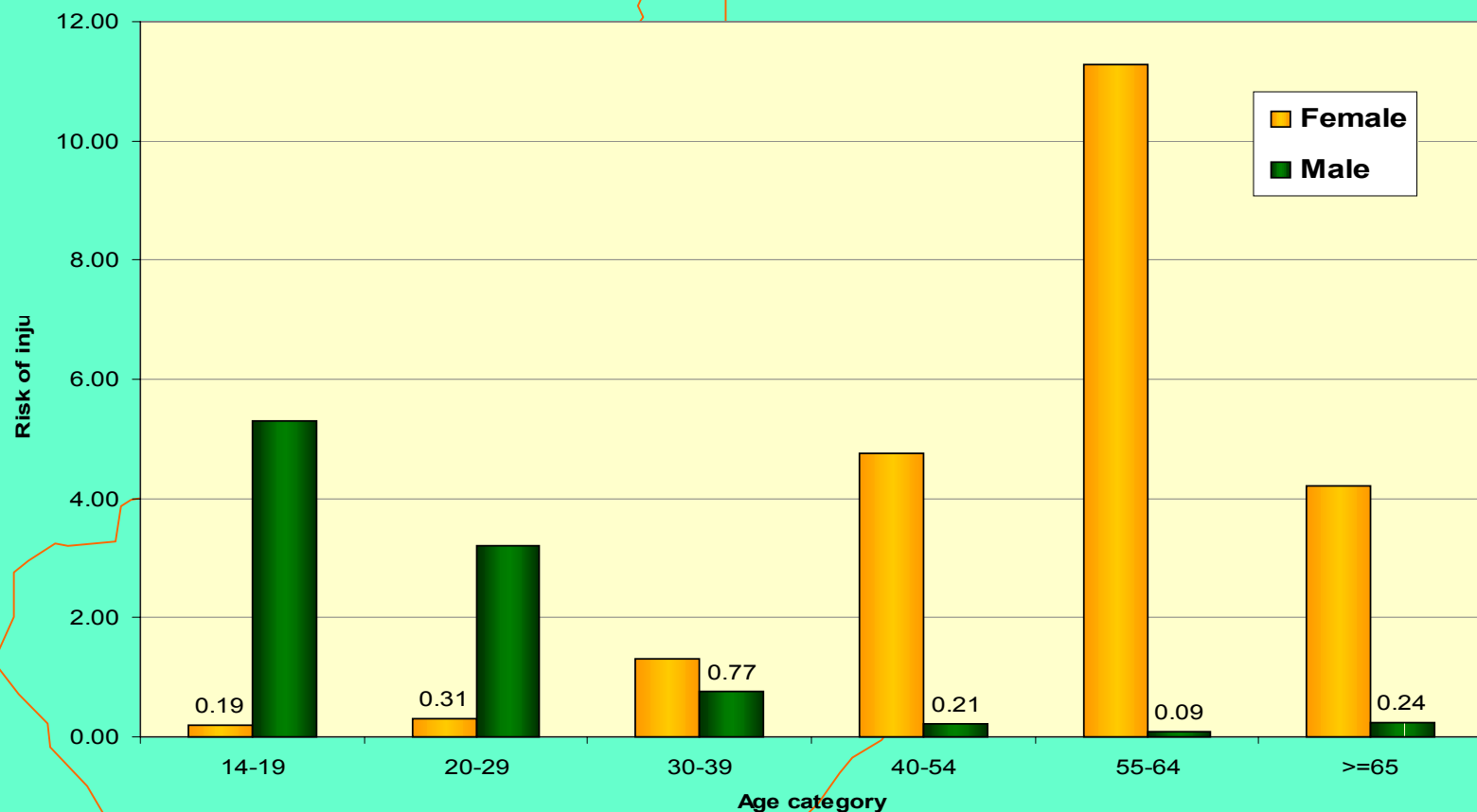
* Risk is based on Proportional Incidence ratios (PIR) relative to the other gender

Gender specific *relative risk of burn injury by age category among cooks over a four-year period

Risk of injury (y-axis)

Legend:
- Female (orange)
- Male (dark green)

Data values by age category:
- 14-19: Male 3.73
- 20-29: Male 2.11
- 30-39: Female 1.97
- 40-54: Female 5.51
- 55-64: Female 15.83
- >=65: Female 3.88

Age category (x-axis): 14-19, 20-29, 30-39, 40-54, 55-64, >=65
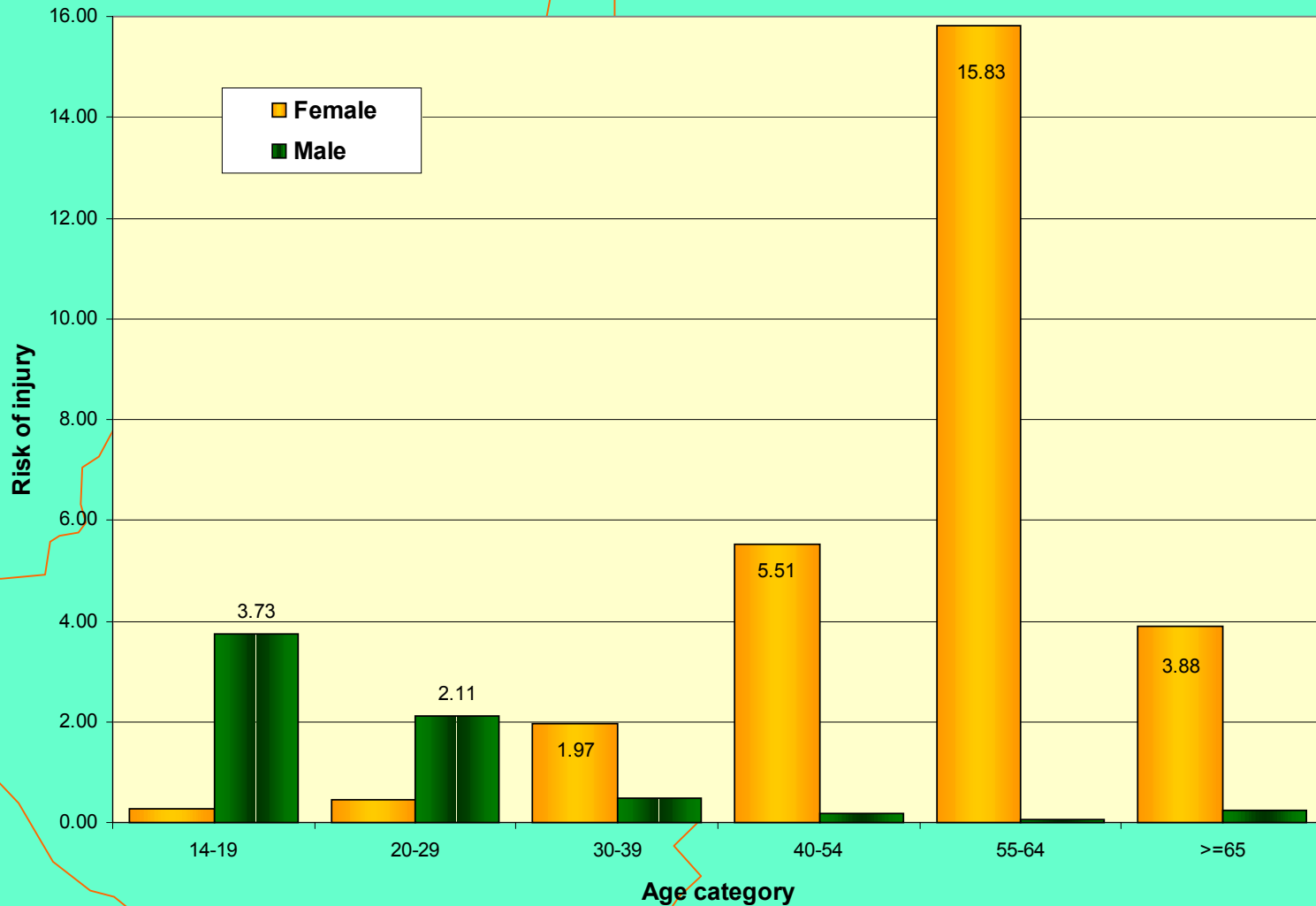
* Risk is based on Proportional Incidence ratios (PIR) relative to the other gender

Distribution of total medical costs for 1995 accident year fracture cases over time

| Period of payment	# unique claims	Average per case cost ($)	Median per case cost ($)	Total cost ($)	Year to year change in total payment
January 1, 1995 to December 31, 1995	2,140	2,098	476	4,489,006	
January 1, 1996 to December 31, 1996	1,948	2,326	440	4,531,699	↑1%
January 1, 1997 to December 31, 1997	733	1,779	353	1,303,753	↓71%
January 1, 1998 to December 31, 1998	361	1,607	481	580,260	↓55%
January 1, 1999 to December 31, 1999	224	1,486	505	332,768	↓43%

Distribution of total medical costs for 1995 accident year back injury cases over time

Period of payment	# unique claims	Average per case cost ($)	Median per case cost ($)	Total cost ($)	Year to year change in total payment
January 1, 1995 to December 31, 1995	8,298	1,109	384	9,201,359	
January 1, 1996 to December 31, 1996	7,364	1830	404	13,475,862	↑46%
January 1, 1997 to December 31, 1997	3,182	2,314	605	7,363,787	↓45%
January 1, 1998 to December 31, 1998	1,923	2,558	776	4,918,258	↓33%
January 1, 1999 to December 31, 1999	1,408	2,485	832	3,489,244	↓29%

Physical Therapy Utilization Reports

These reports contain counts of events defined by payments made for both defined appropriate and inappropriate physical therapy treatments as identified by **cpt** codes in the existing treatment guidelines . Both reports contain counts of physical therapy treatment events during the **first 30 days**, **the second 30 days**, **the first year after the initial 60 days** and **the second year after the initial 60 days.**

Each report also provides the opportunity to view the detail payments for each claim number.
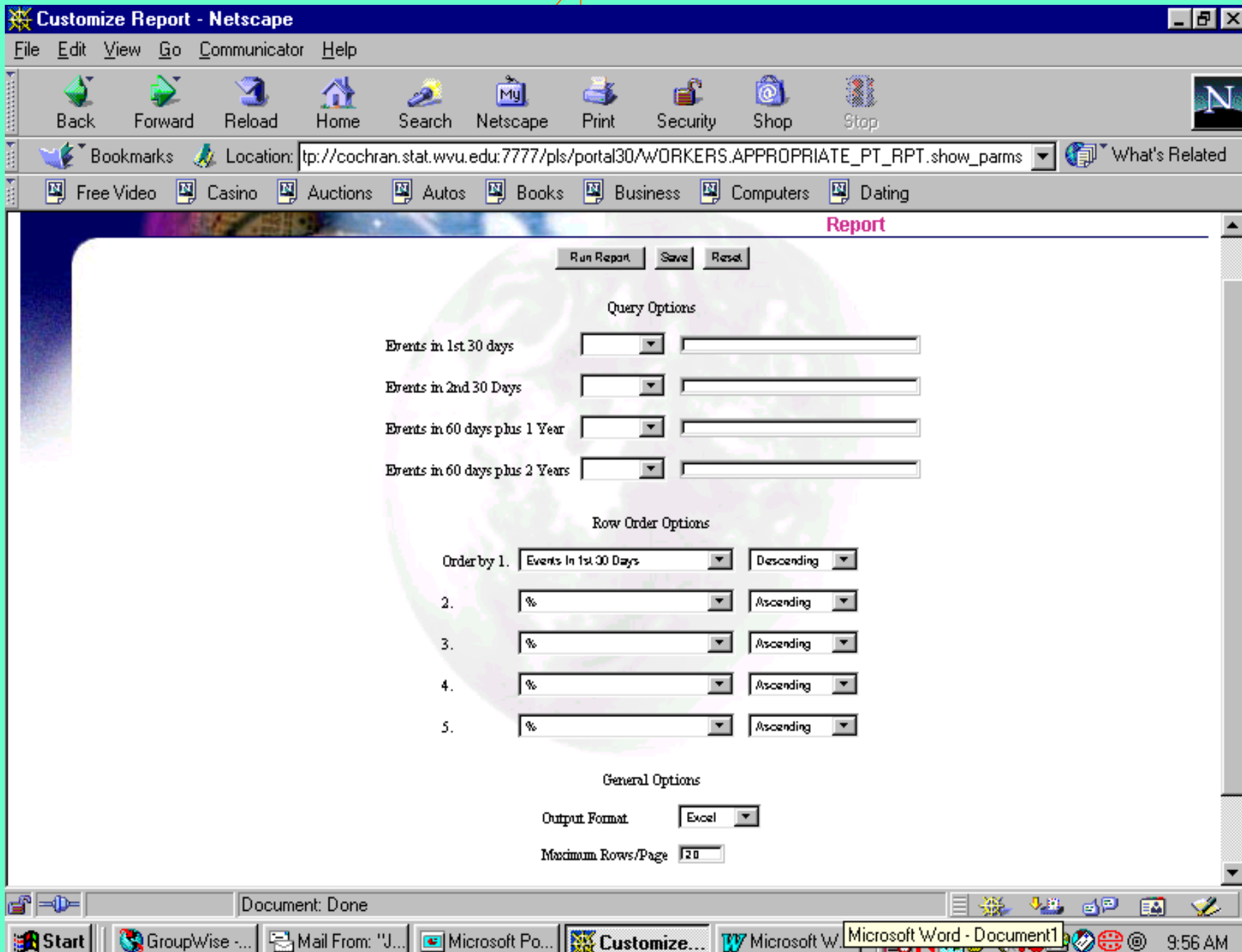
Defined Appropriate Physical Therapy Utilization

Claim Num	Events In 1st 30 Days	Events In 2nd 30 Days	Events In 60 Days And 1 Year	Events In 60 Days And 2nd Year
990014700	29	3	34	0
990051859	28	9	0	0
960002760	28	20	19	24
960005544	27	21	59	0
980030912	27	24	130	62
2001045364	27	15	0	0
970062904	26	14	85	26
980058253	26	19	88	0
960005322	26	23	183	65
960013186	26	20	23	4
950052554	25	8	25	30
960009450	25	19	11	0
970003123	25	22	8	0
970016480	25	1	24	0
970045643	25	9	0	0
980020236	25	14	23	7
980045183	25	0	0	0
980059635	25	0	18	0
990034466	25	19	40	1
990055277	24	12	11	0

Next

Defined Appropriate Physical Therapy

Claim Num	Service Date	Service Allow Amt	Vendor Num
990014700	July 29, 1998	$28.77	1021024
	July 30, 1998	$28.77	1021024
	July 31, 1998	$28.77	1021024
	August 2, 1998	$28.77	1021024
		$28.77	1021024
	August 3, 1998	$28.77	1021024
		$28.77	1021024
	August 4, 1998	$28.77	1021024
		$28.77	1021024
	August 5, 1998	$28.77	1021024
		$28.77	1021024
	August 6, 1998	$28.77	1021024
		$28.77	1021024
	August 7, 1998	$28.77	1021024
		$28.77	1021024
	August 8, 1998	$28.77	1021024
		$28.77	1021024
	August 9, 1998	$28.77	1021024
		$28.77	1021024
	August 10, 1998	$28.77	1021024

Next

Conclusions

Acknowledging both the limitations and complexity of large insurance data sets, we still see several important Public Health uses:

1. Surveillance for injury and Illness outbreaks- Creation of data cubes containing pre-aggregated facts or outcomes makes it feasible to use large inefficient insurance data in Public Health Surveillance.

2. Analysis of Service and outcome Trends - As the data cubes can be updated frequently, the information is timely and accurate. Underneath the data cube lies the original data which can be analyzed for transactional questions such as number of physical therapy visits provided in first 30 days and whether or not they were timely or exceeded the recommended guidelines. Therefore, appropriateness of a particular service is based on definition and not on technology or database.

3. Active and Pro-Active monitoring- The flexibility of the system is that it allows both active and proactive monitoring of facts (outcomes). Active monitoring is based on WEB-EIS and proactive monitoring is based on detail data. So, using WEB-EIS we can monitor unusual trends while using detailed data we can identify unusual cases in terms of occurrence, utilization and cost.

4. Applicability to a variety of settings- Once developed, this system will provide powerful tool to the monitors of standard of care (i.e., HbA1c, ophthalmology and podiatry visit for diabetes).